



ASA-1154

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of

Y. HIRAIWA et al

Serial No. 10/649,671

Group Art Unit: 2186

Filed: August 28, 2003

Examiner: Unassigned

For: DISK SYSTEM, COMPUTER SYSTEM, STORAGE MANAGING METHOD AND PROGRAM

TRANSMITTAL OF CERTIFIED PRIORITY DOCUMENT

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Submitted herewith is a certified priority document (JP 2002-320957) of a corresponding Japanese patent application for the purpose of claiming foreign priority under 35 U.S.C. § 119. An indication that this document has been safely received would be appreciated.

Respectfully submitted,

John R. Mattingly
Registration No. 30,293
Attorney for Applicants

MATTINGLY, STANGER & MALUR
1800 Diagonal Rd., Suite 370
Alexandria, Virginia 22314
(703) 684-1120
Date: June 2, 2004

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2002年11月 5日
Date of Application:

出願番号 特願2002-320957
Application Number:
[ST. 10/C]: [JP2002-320957]

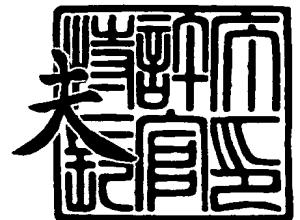
出願人 株式会社日立製作所
Applicant(s):

USPN 10/649,671
MATTINGLY, STANGER + MACUR
(703) 684-1120
DKT: ASA-1154

2003年 8月26日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



出証番号 出証特2003-3069623

【書類名】 特許願

【整理番号】 NT02P0683

【提出日】 平成14年11月 5日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 12/00501

G06F 3/06

G06F 17/60

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 平岩 友理

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 長須賀 弘文

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 清井 雅広

【発明者】

【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

【氏名】 宮田 和久

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 櫻庭 健年

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100068504

【弁理士】

【氏名又は名称】 小川 勝男

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100086656

【弁理士】

【氏名又は名称】 田中 恭助

【電話番号】 03-3661-0071

【選任した代理人】

【識別番号】 100094352

【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ディスクサブシステム、計算機システム、それを管理するためのストレージ管理方法、および、管理プログラム

【特許請求の範囲】

【請求項 1】 ネットワークを介して計算機から論理ボリュームとしてアクセスされるディスクストレージを備えるディスクサブシステムにおいて、

このディスクストレージは、

CPUと、

記憶装置とを有し、

前記CPUにより、前記ディスクストレージの有するボリューム管理情報を読み込んで、論理ボリューム構成情報を構成し、前記記憶装置に、その論理ボリューム構成情報を保持することを特徴とするディスクサブシステム。

【請求項 2】 計算機と、ネットワークを介して前記計算機から論理ボリュームとしてアクセスされるディスクストレージを備えるディスクサブシステムとを有する計算機システムにおいて、

さらに、この計算機システムは、管理計算機を有し、

前記ディスクストレージは、

CPUと、

記憶装置とを有し、

前記CPUにより、前記ディスクストレージの有するボリューム管理情報を読み込んで、論理ボリューム構成情報を構成し、前記記憶装置に、その論理ボリューム構成情報を保持して、

前記管理計算機から、その論理ボリューム構成情報に基づくボリューム管理情報抽出データをアクセスできるようにしたことを特徴とする計算機システム。

【請求項 3】 ネットワークを介して計算機から論理ボリュームとしてアクセスされるディスクストレージを備えるディスクサブシステムのストレージ管理方法において、

(1) 前記ディスクストレージの有するボリューム管理情報を読み込むステップ

(2) 前記管理情報から前記論理ボリューム構成情報を構成するステップを有し

、
前記(1)(2)のステップをこの順に実行することを特徴とするストレージ管理方法。

【請求項4】 さらに、

(3) 前記論理ボリューム構成情報に基づくボリューム管理情報抽出データを作成して、前記ネットワークに接続されている管理計算機に送信するステップを有し、

前記(1)(2)のステップの実行の後に、前記(3)のステップを実行することを特徴とする請求項3記載のストレージ管理方法。

【請求項5】 ネットワークを介して計算機から論理ボリュームとしてアクセスされるディスクストレージを備えるディスクサブシステムのボリューム管理のための管理プログラムにおいて、

前記ディスクストレージの有するボリューム管理情報を読み込む機能、

前記管理情報から前記論理ボリューム構成情報を構成する機能、

前記論理ボリューム構成情報に基づくボリューム管理情報抽出データを作成して、前記ネットワークに接続されている管理計算機に送信する機能、

前記管理計算機に、前記ボリューム管理情報抽出データを出力する機能を実現するための管理プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、ストレージ管理方法に係り、特に、管理情報の収集にあたり、ホスト側計算機とネットワークに負荷をかけることの少ないストレージ管理方法に関する。

【0002】

【従来の技術】

近年の社会基盤を担う公共・企業の基幹システムにおけるデータ量増加に対応するため、ストレージシステムにおいては、ディスクアレイの制御技術を用いて

ストレージの大容量化に対応してきた。特許文献1には、このようなディスクアレイを用いる技術が開示されている。

【0003】

また、ストレージ専用のネットワークであるSAN (Storage Area Network) を用いてこれらの基幹システムを構築し、ストレージの大容量化および複雑化に対応する傾向が現れている。このようなシステムにおいては、サーバとストレージの間をSANで接続し、ストレージをサーバから独立管理させて柔軟な管理や拡張をおこなうことが望まれている。特許文献2には、このようなSANを用いたシステムの管理方法についての技術が公開されている。この技術により、RAID (Redundant Arrays of Inexpensive Disks) や、MT (Magnetic Tape) 等のストレージデバイスを、一元運用することができるようになる。

【0004】

また、システムに複数のサーバがあり、それぞれに複数のディスクデバイスが割り当てられているSAN環境における負荷分散を目指した運用方法に、特許文献3がある。これは、ディスクに対するアクセス負荷分散を図るとともに、そのためのリソース監視をおこなう技術を開示したものである。

【0005】

【特許文献1】

特開2001-256003号公報

【特許文献2】

特開2002-7304号公報

【特許文献3】

特開2002-175203号公報

【0006】

【発明が解決しようとする課題】

上記特許文献3では、業務を行うサーバから、リソース管理をおこなう管理サーバに情報を収集する際に、業務をおこなうサーバにリソース監視をおこなうプログラムを設けて、そこから定期的にディスクボリュームの負荷情報を収集していた。したがって、定期的に負荷収集のためのプログラムがCPUを使用するこ

とになり、業務をおこなうためのCPUに負荷がかかっていた。

【0007】

また、SAN環境での一元運用のために必要な情報として、特許文献3に記載されているアクセス情報に加え、ディスクの空き情報等が挙げられる。この情報を取得するためにサーバにリソース監視をおこなうプログラムを置く方法をとった場合には、サーバ側からディスクに対して空き情報等の情報を取得する命令を発行しなければならない。このため、業務をおこなうためのI/O命令を遅延させる可能性があった。

【0008】

本発明は、上記問題点を解決するためになされたもので、その目的は、ディスクサブシステムのストレージの管理情報の収集にあたり、ホスト計算機とネットワークの負荷を軽減して、効率的なストレージアクセスをおこなうことのできる計算機システムを提供することにある。

【0009】

【課題を解決するための手段】

本発明は、CPUを具備したディスクサブシステムと、それを参照する計算機と、システムを管理するための管理計算機を具備し、各々が相互にネットワークで接続された計算機システムにおいて、管理計算機からの要求に従ってディスクの情報を収集し、報告する情報収集プログラム（エージェント）を、上記CPUを具備したディスクサブシステム上で動作させるものである。

【0010】

本発明の計算機システムによれば、エージェントをディスクサブシステム上で動作させるため、業務をおこなう計算機に負荷をかけることなく、ディスクの情報を収集することができる。また、ディスクの情報を収集する際に、計算機を介することなく管理計算機に伝送するため、計算機とディスクの間の通信路に負荷をかけることなく、ディスクの情報を収集することができる。

【0011】

【発明の実施の形態】

以下、本発明に係る各実施形態を、図1ないし図6を用いて説明する。

【0012】

本発明は、ストレージ管理方法に関するものである。ストレージ管理のためには、さまざまなデータを扱うことが考えられるが、本実施形態では、管理プログラムが必要とするディスクボリュームのデータは、ディスクの空き容量、総容量、アクセス頻度の三つであるものとして説明する。

【0013】

〔計算機システムの構成〕

先ず、図1を用いて本発明に係る計算機システムの構成を説明する。

図1は、本発明に係る計算機システムの構成図である。

【0014】

本システムは、図1に示されるように、管理計算機100と、管理対象の計算機800と、計算機からアクセスされるディスクサブシステム900で構成され、これらがネットワークで接続されている。ディスクサブシステム900には、ディスクサブシステム900内のディスク管理をおこなうためのSVP (Service Processor) 1000が付属している。計算機800とディスクサブシステム900は、ストレージ専用のネットワークであるSAN 1100で接続されている。また、管理計算機100と計算機800とSVP 1000は、LAN (Local Area Network) 1200で接続されている。

【0015】

ユーザの業務をおこなうAP (Application Program) 803は、計算機800の主記憶装置802上にロードされ、CPU 801で実行される。AP 803は、業務を処理するため、必要に応じてディスクサブシステム900に対して、ファイル毎のデータの読み出しおよび書き込み (アクセス) をおこなう。

【0016】

ディスクサブシステム900には、ディスクコントローラ903によって制御されるディスク904がある。なお、ディスク904は仮想的なディスク (論理ボリューム) であってもよく、各々が二つ以上のディスクから構成されたアレイであってもよいが、本実施形態では、ディスク904に対するI/O命令は、ディスクIDで識別され、かつ、ディスク単位で論理ボリュームとして各計算機8

00に割り当てられており、ディスク単位で情報の収集・管理をおこなうものとする。

【0017】

ディスク904上には、割り当てられている計算機800のファイルシステムに依存したフォーマットのボリューム管理情報200が置かれている。

【0018】

ディスクサブシステム900の主記憶装置902上には、リソース監視をおこなうエージェント700が展開され、CPU901で実行される。エージェント700は、ディスクサブシステム上におかれるボリュームの管理情報収集のためのプログラムであり、管理情報を出力するボリュームを指定したボリューム指定情報300と、そのボリュームの参照方法を記述したボリューム構成情報400と、監視プログラム600から構成される。監視プログラム600は、ボリューム指定情報300とボリューム構成情報400の二つの情報に基づき、SVP1000の主記憶装置1002上に設けたボリューム管理情報抽出データ500に管理情報を出力する。そして、管理計算機100の要請があったときには、LAN1200を介して、このボリューム管理情報抽出データ500を送信する。

【0019】

エージェント700は、SVP1000が有する補助記憶装置1003に格納されていて、必要に応じてSVP1000からディスクサブシステム900の主記憶装置902上に展開され、実行される。

【0020】

また、展開後もボリューム指定情報300を、SVP1000から更新することとで、ディスク904の構成変更にも対応することができる。また情報取得をおこなうか否かを変更することもできる。

【0021】

管理計算機100の主記憶装置102上には、管理プログラム103が展開されており、CPU101で実行される。この管理プログラム103が、SVP1000のボリューム管理情報抽出データ500を読み込んで、ディスクサブシステム900の情報を収集する。

【0022】

〔ストレージ管理のためのデータ構造〕

次に、図2ないし図5を用いて本発明のストレージ管理のためのデータ構造を説明する。

図2は、ボリューム管理情報200とボリュームの構成図である。

図3は、ボリューム指定情報300の模式図である。

図4は、ボリューム構成情報400の模式図である。

図5は、ボリューム管理情報抽出データ500の模式図である。

【0023】

ボリューム管理情報200は、そのボリュームの管理のためのデータであり、図2に示されるようにディスク904上におかれる。

【0024】

ディスク904は、ボリューム管理情報200の開始位置を示す先頭位置209を有している。ディスクID201は、そのボリュームを識別するIDを格納する領域である。空き容量202は、そのボリュームの空き容量を格納した領域である。総容量203は、そのボリュームの総容量を格納した領域である。先頭コントロールブロックへのポインタ204は、ファイルで使用する領域を管理するためのコントロールブロック205のリストの先頭へのポインタを格納した領域である。

【0025】

コントロールブロック205は、そのボリュームの中で、ファイルが使用するデータブロックを管理するためのリスト状の構造を持つテーブルであり、ファイル情報206、次コントロールブロックへのポインタ207、データブロックへのポインタ208から構成される。ファイル情報206は、どのファイルがそのデータブロックを使用しているかを識別する情報を格納する領域である。空きであることを示す場合もある。次コントロールブロックへのポインタ207は、次のコントロールブロック205へのポインタを格納する領域である。データブロックへのポインタ208は、ボリュームの中のデータブロックへのポインタを格納する領域である。

【0026】

ボリューム指定情報300は、ボリュームの管理情報を収集するための情報を記述したものであり、図3に示されるように、ディスクID301、情報収集要否302、ボリューム構成ID303からなる。

【0027】

ディスクID301は、ボリュームを識別するIDを格納する領域である。情報収集要否302は、そのボリュームの情報収集をおこなうか否かの指定を格納する領域であり、この領域の値により、監視プログラムは、ボリュームの管理情報を収集するか否かを決定する。

【0028】

本実施形態では、収集をおこなう場合は「1」が、おこなわない場合は「0」が記述されているものとする。ボリューム構成ID303は、そのボリュームのボリューム管理情報904から情報を読むために、ボリューム構成情報400のどのIDの情報をを用いるかを示した情報を格納した領域である。

【0029】

ボリューム構成情報は、ボリューム管理のために必要な情報を抜き出したものであり、図4に示されるように、ボリューム構成ID401、先頭位置情報402、空き情報記述位置403からなる。

【0030】

ボリューム構成ID401は、ボリューム構成の種類を識別するIDを格納した領域であり、図3のボリューム指定情報のボリューム構成IDとリンクしている。

【0031】

先頭位置情報402は、ボリューム管理情報200の先頭位置を格納した領域である先頭位置206が、ボリューム内のどこに置かれるかの情報が記述される。空き容量記述位置403は、ボリューム管理情報200の中の空き容量202の位置を格納した領域である。総容量記述位置404は、ボリューム管理情報200の中の総容量203の位置を格納した領域である。

【0032】

なお、本実施形態の記法で、空き情報記述位置 403 と総容量記述位置 404 が、括弧でくくられて記述されている場合は、先頭位置情報 402 から得られたボリューム管理情報 200 の先頭からの偏差を表し、括弧がない場合は、ボリューム内での位置を示しているものとして説明する。

【0033】

ボリューム管理情報抽出データ 500 は、必要なボリュームの管理情報を SVP1000 で保持する情報であり、図 5 に示されるように、ディスク ID 501、空き容量 502、総容量 503、アクセス頻度 504 からなる。

【0034】

ディスク ID 501 は、ボリュームを識別する ID を格納する領域である。空き容量 502 は、監視プログラム 600 が取得したディスク ID 501 の空き容量を格納する領域である。総容量 503 は、監視プログラム 600 が取得したディスク ID 501 の総容量を格納する領域である。アクセス頻度 504 は、監視プログラム 600 が取得したディスク ID 501 のボリュームのアクセス頻度を格納する領域である。

【0035】

〔監視プログラムの処理〕

次に、図 6 を用いて監視プログラム 600 の処理の流れを説明する。

図 6 は、監視プログラム 600 の処理の流れを示すフローチャートである。

【0036】

まず、監視プログラムは、ボリューム指定情報 300 から、情報収集要否 302 が「1」であるディスク ID 301 を取得する（ステップ 601）。

【0037】

そして、ステップ 601 で取得したディスク ID 301 が空でないか、すなわち、情報を収集するディスクがあるか否かを判定する（ステップ 602）。

【0038】

情報を収集するディスクがある場合には、ディスク ID 301 に対応するボリューム構成 ID 303 を取得し（ステップ 603）、情報を収集するディスクがない場合には、監視プログラムの一回分の処理を終了したとして、次の起動契機

までWAITする（ステップ608）。

【0039】

情報を収集するディスクがある場合で、ステップ603が実行されたときは、次に、ディスクID301で示されるディスク904のボリューム管理情報200から、情報を取得する。すなわち、ボリューム構成情報400の中から、ステップ603で取得したボリューム構成ID303に合致するボリューム構成ID401を見つけ、それに対応する先頭位置情報402と、空き情報記述位置403と、総容量記述位置404を取得し、この二つの情報を用いて、ボリューム管理情報200から、空き容量202と総容量203の値を取得する。

【0040】

例えば、ディスクID301が「1」である場合は、ボリューム構成ID401の値である「1」に対応する先頭位置情報402は「4バイト」である。また、空き情報記述位置403は「（4バイト）」、総容量記述位置404は「（8バイト）」である。したがって、ディスクID301が「1」であるディスク904の先頭から4バイトに記述されているポインタが示す位置（ボリューム管理情報200の開始位置）から、さらに4バイト偏移した位置から空き容量202を取得し、また、同じくボリューム管理情報200の開始位置から8バイト偏移した位置から総容量203を取得する（ステップ604）。

【0041】

さらに、ディスクID301で示されるディスク904のアクセス頻度情報を取得する（ステップ605）。

【0042】

ステップ604とステップ605で取得した情報を、ボリューム管理情報抽出データ500に出力する（ステップ606）。

【0043】

ボリューム指定情報300から、次に処理すべきディスクのディスクIDを取得する（ステップ607）。そして、このようにして処理するディスクIDの全てについて、ステップ602ないしステップ607の処理を繰り返す。

【0044】

〔その他の実施形態〕

上記実施形態においては、ディスクの中に格納されているボリューム管理情報から、管理プログラムが必要とするデータを抽出する形態を採っていた。これを、ボリューム管理情報の始点と終点を指定させ、ボリューム管理情報全体を出力する形態にしてもよい。この場合、ボリューム管理情報抽出データ 500 には、ボリューム管理情報 200 全体が格納される。

【0045】

そして、このときには、ボリューム管理情報全体を出力するのではなく、前回出力分からの差分を出力する形態にしてもよい。

【0046】

また、上記実施形態の管理情報としては、ボリュームの空き容量、総容量、アクセス頻度の三種類のものを示したが、そのほかにも、ボリュームの計算機の割り当て情報、RAIDのエミュレーションタイプ、ディスク単位のI/O書込みミス頻度などが考えられる。

【0047】**【発明の効果】**

本発明によれば、ディスクサブシステムのストレージの管理情報の収集にあたり、ホスト計算機とネットワークの負荷を軽減して、効率的なストレージアクセスをおこなうことのできる計算機システムを提供することができる。

【図面の簡単な説明】**【図1】**

本発明に係る計算機システムの構成図である。

【図2】

ボリューム管理情報 200 とボリュームの構成図である。

【図3】

ボリューム指定情報 300 の模式図である。

【図4】

ボリューム構成情報 400 の模式図である。

【図5】

ボリューム管理情報抽出データ 5 0 0 の模式図である。

【図 6】

監視プログラム 6 0 0 の処理の流れを示すフローチャートである。

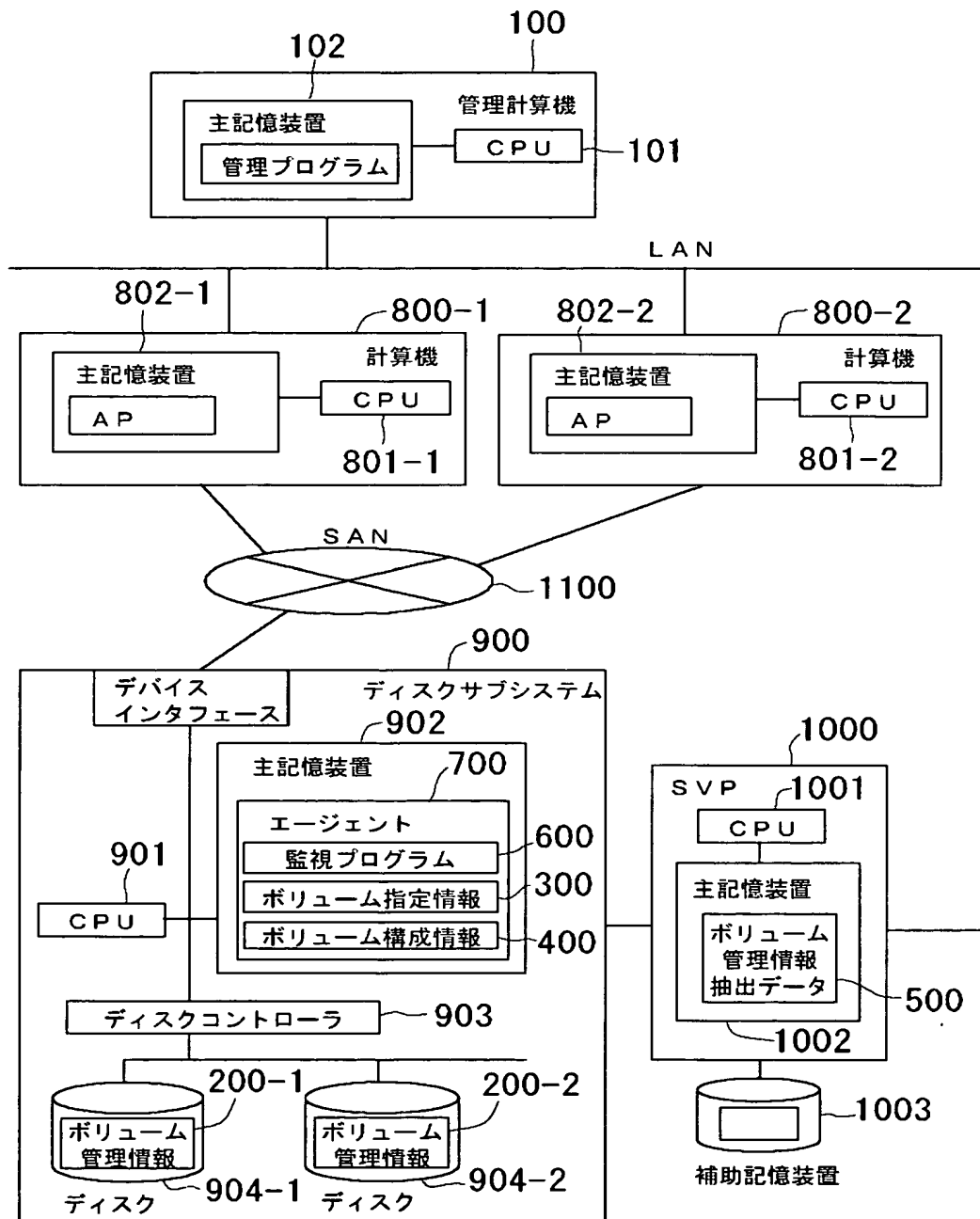
【符号の説明】

1 0 0 … 管理計算機、1 0 3 … 管理プログラム、2 0 0 … ボリューム管理情報、3 0 0 … ボリューム指定情報、4 0 0 … ボリューム構成情報、5 0 0 … ボリューム管理情報抽出データ、6 0 0 … 監視プログラム、7 0 0 … エージェント、8 0 0 … 計算機、9 0 0 … ディスクサブシステム、1 0 0 0 … SVP、1 1 0 0 … SAN、1 2 0 0 … LAN。

【書類名】 図面

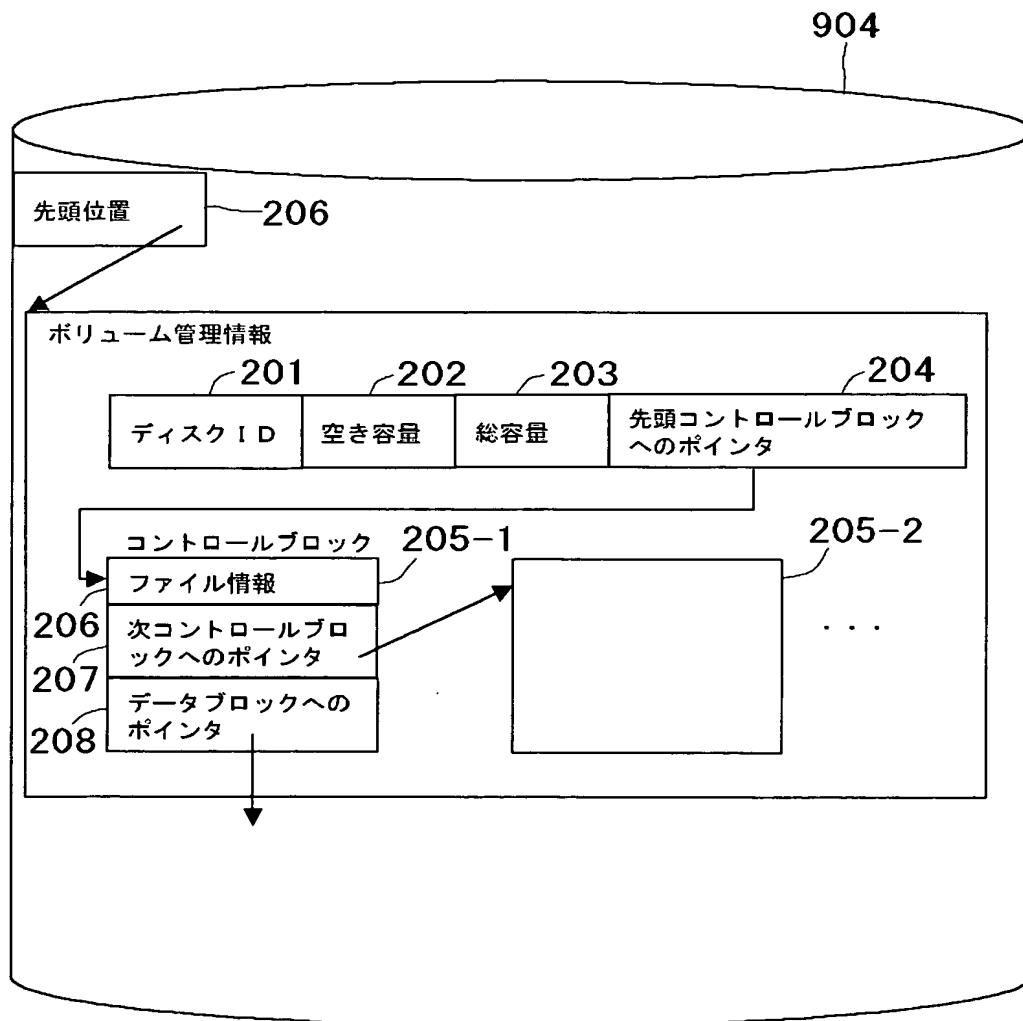
【図 1】

図 1



【図 2】

図 2



【図 3】

図 3

300

ボリューム指定情報		
ディスクID	情報収集要否	ボリューム構成ID
1	1	1
2	0	5

301 302 303

【図 4】

図 4

400

ボリューム構成情報			
ボリューム構成 ID	先頭位置情報	空き容量記述位置	総容量記述位置
1	4 バイト	(4 バイト)	(8 バイト)
5	16 バイト	18 バイト	24 バイト

401
402
403
404

【図 5】

図 5

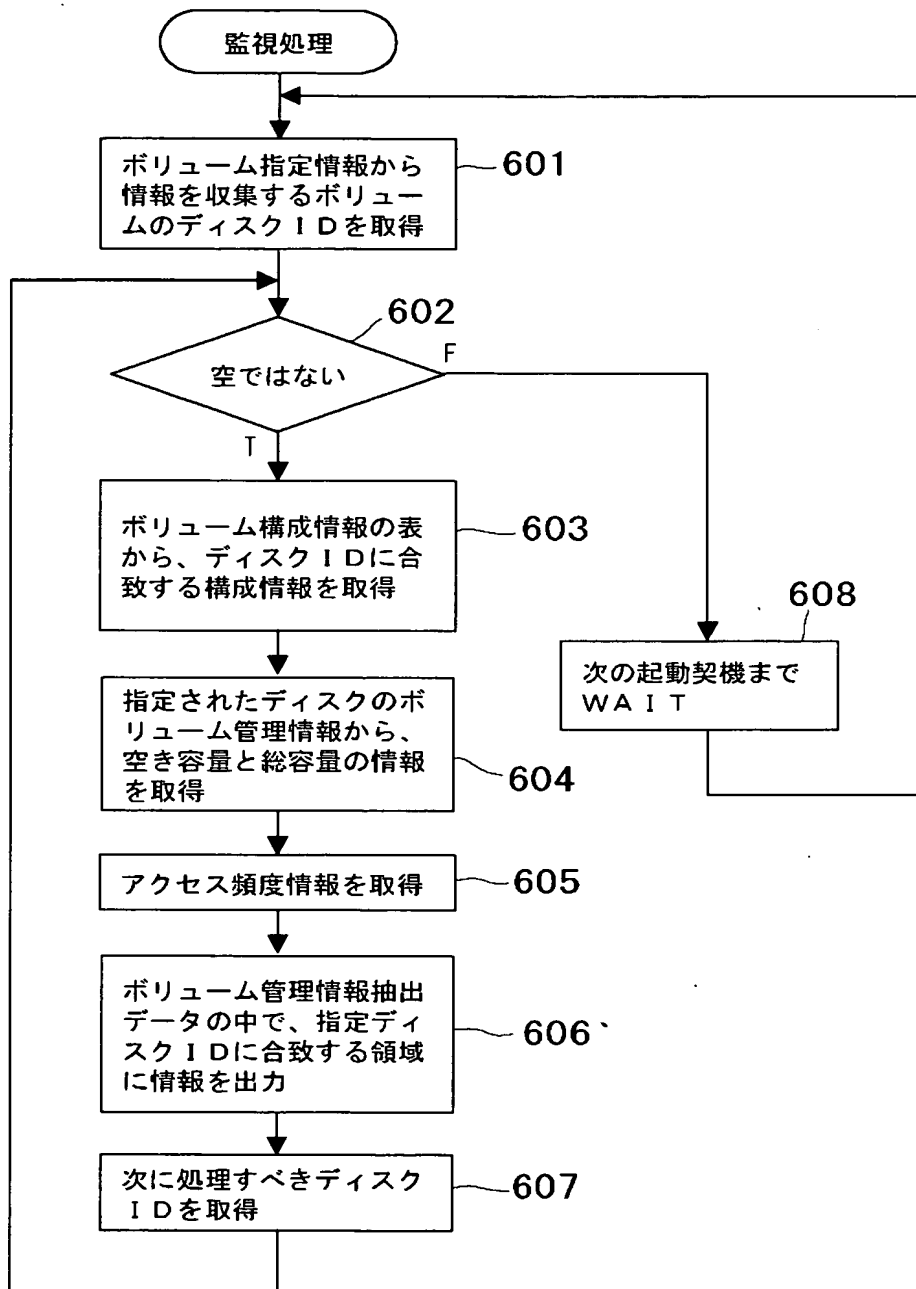
500

ボリューム管理情報抽出データ			
ディスクID	空き容量	総容量	アクセス頻度
1			

501502503504

【図 6】

図 6



【書類名】 要約書

【要約】

【課題】 ディスクサブシステムのストレージの管理情報の収集にあたり、ホスト計算機とネットワークの負荷を軽減して、ストレージを有する計算機システムのスループットを向上させる。

【解決手段】 ディスクサブシステムに、CPUと主記憶装置を設け、管理計算機からの要求にしたがってディスクの情報を収集する。情報収集プログラムを、ディスクサブシステム上で動作させる。これにより、ディスクサブシステム上でボリューム管理情報を編集し、管理計算機で利用できるようにする。

【選択図】 図 1

特願 2 0 0 2 - 3 2 0 9 5 7

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1. 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所